Computers in Human Behavior 54 (2016) 301–309

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

The relationship between metacognitive experiences and learning: Is there a difference between digital and non-digital study media?

Elisabeth Norman*, Bjarte Furnes

Faculty of Psychology, University of Bergen, Norway

ARTICLE INFO

Article history:
Received 10 December 2014
Received in revised form
18 June 2015
Accepted 22 July 2015
Available online 27 August 2015

Keywords:
Metacognition
Metacognitive experiences
Digital learning
Reading
Learning outcome
Judgement of learning
Prediction of performance
Confidence rating

ABSTRACT

Technological development has influenced the ways in which learning and reading takes place, and a variety of technological tools now supplement and partly replace paper books. Previous studies have suggested that digital study media impair metacognitive monitoring and regulation (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman 2014). The aim of the current study was to explore the relationship between metacognitive experiences and learning for digital versus non-digital texts in a test situation where metacognitive experiences were assessed more broadly compared to previous studies, and where a larger number of potentially confounding factors were controlled for. Experiment 1 ($N = 100$) addressed the extent to which metacognitive monitoring accuracy for 4 factual texts was influenced by whether texts were presented on a paper sheet, a PC, an iPad, or a Kindle. Metacognitive experiences were measured by Predictions of Performance (PoP), Judgements of Learning (JoL), and Confidence Ratings (CR), and learning outcome was measured by recognition performance. Experiment 2 ($N = 50$) applied the same basic procedure, comparing a paper condition with a PC condition with the opportunity to take notes and highlight text. In both experiments, study media had no consistent effect on metacognitive calibration or resolution. The results give little support to previous claims that digital learning impairs metacognitive regulation.

1. Introduction

In today's society there is an increased use of digital equipment, with PC's and tablet devices now being used more frequently also in educational settings. This has opened up for new ways of learning, both at an individual level but also at a group level. For instance, there is currently a large interest in the development of collaborative e-learning environments and multidisciplinary learning groups (e.g., Dascalua, Bodea, Lytras, Ordoñez de Pablos, & Burlacua, 2014), and technology is also seen as an important element of knowledge management (e.g., Zhao & Ordóñez de Pablos, 2011). This development calls for more knowledge about if and to what extent cognition is influenced by digital versus non-digital presentation format (Carr, 2010). In an educational context, digitalization has resulted in an increased emphasis on students' *digital competence*. In parallel, there is an additional focus in today's schools on students' ability to engage in *self-regulation*, defined as the extent to which the learner is “metacognitively, motivationally and behaviourally active participants in their own learning process” (Zimmerman, 1986, p. 308). The combined focus on digital competence and self-regulation necessitates more knowledge about the relationship between learning and self-regulation in digital compared to traditional paper-based learning.

According to the above definition, self-regulation refers to students' ability to regulate learning through metacognitive processes. From this perspective, self-regulation requires the ability to engage in *metacognition*, i.e., cognition about one's own cognition (Metcalfe, 2000). A distinction is often made between *metacognitive monitoring*, where metacognitive thoughts or feelings reflect aspects of ongoing cognitive processing, and *metacognitive control*, where the output of such monitoring is used to regulate cognitive processes and behaviour (Koriat, 2000, 2007). An example would be the decision to read a text once more if one felt that comprehension was low. Within such a framework, metacognitive monitoring is a prerequisite for metacognitive control and self-regulation.

Self-regulation and metacognition have become central concepts in a wide variety of studies on online learning, e-learning and digital media use. For instance, a recent study by Pellas (2014) that

* Corresponding author. Psychology Faculty, University of Bergen, Christies gate 12, 5020 Bergen, Norway.
E-mail address: Elisabeth.Norman@uib.no (E. Norman).

http://dx.doi.org/10.1016/j.chb.2015.07.043

explored student engagement and learning in a virtual reality learning environment, found that metacognitive self-regulation was one predictor of emotional and cognitive engagement. Similarly, in an experimental study looking at "preflective prompts" (i.e., a request for reflection before the learning task) in e-learning, Lehmann, Hähnlein, and Ifenthaler (2014) found that metacognitive awareness was a significant predictor of effective self-regulation.

In this paper we focus on the relationship between learning and metacognitive monitoring in digital and non-digital learning contexts where the learning material is written texts. Previous research has shown that people's subjective preferences for reading texts are often in favour of paper-based rather than digital formats (Buzzetto-More, Sweat-Guy, & Elobaid, 2007; Jamali, Nicholas, & Rowlands, 2009; Spencer, 2006; Woody, Daniel, & Baker, 2010). A number of studies have also explored effects of study media on learning outcome (e.g., Mangen, Waldermo, & Brønnick, 2013) and on subjective experiences (e.g., Mangen, 2006), of which metacognitive experiences would be one example. However, to our knowledge only three studies to date have specifically compared how learning and metacognition is related in digital versus non-digital learning contexts (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014). Here the relationship between learning and metacognitive monitoring was measured as the degree of correspondence between memory performance and *prediction of performance* (PoP), reported either after the study participants had completed an entire text or at regular time intervals during text reading. Absolute monitoring accuracy, referred to as *calibration bias*, was calculated as the absolute difference between memory and total PoP. One of the studies (Ackerman & Goldsmith, 2011) also included a measure of relative monitoring accuracy, referred to as *metacognitive resolution*, which is the correlation between PoP's and recognition scores for a series of texts. The general finding in these studies was that participants in on-screen conditions showed more overconfidence than participants in on-paper conditions, interpreted as calibration bias being influenced by study media. This was found for both free and fixed study time. The only study that included a measure of metacognitive resolution, i.e., relative monitoring accuracy, found that this was not influenced by study media (Ackerman & Goldsmith, 2011). As to the question of whether study media influences learning outcome, results from the studies were mixed.

If metacognitive monitoring and regulation are influenced by presentation format, this has potentially wide-ranging implications both for teaching and research. For instance, it could mean that educators should adjust their expectations of student performance depending on whether a test is conducted on screen or on paper, and also address how students' metacognitive skills in digital learning contexts can be improved. Furthermore, it may encourage researchers to include study media as a potentially relevant variable in research on study processes and metacomprehension. However, in our opinion there is reason to be cautious about drawing such inferences on the basis of the aforementioned studies alone. One reason is the relatively small total number of studies and participants, and the need to replicate the basic effect. Another reason is some potential shortcomings of the basic paradigm used. In the following, we outline each of these, and present two experiments that were specifically designed to address these concerns.

1.1. Measuring metacognitive experiences

One potential shortcoming of the above studies is that they only include one measure of metacognitive experiences, namely PoP. Because reading involves a wide range of cognitive activities, it is likely that a variety of different forms of metacognitive experiences may arise in conjunction with these activities both before, during, and after reading. In order to better capture possible differences in metacognitive experiences across study media and thereby increase the validity of the findings, one should therefore broaden the range of metacognitive measures applied.

A related point is that only one of the studies measured metacognitive resolution. Whereas calibration bias refers to the person's ability to estimate their actual performance level, metacognitive resolution refers to the ability to discriminate between differences in memorability of individual knowledge units (Dunlosky & Metcalfe, 2009). The only measure of metacognitive resolution included in the study by Ackerman and Goldsmith (2011) was also based on PoP. Because PoP was measured either once for each text (Experiment 1) or every 5 min during reading (Experiment 2), each individual correlation was based on very few data points. This statistical limitation was also pointed out by the authors, who referred to recent criticisms of the use of gamma correlations in metacognition research (Benjamin & Diaz, 2008; Masson & Rotello, 2009). Moreover, without specifically controlling for which part of the text each PoP refers to, the degree of correspondence between PoP and performance does not necessarily reflect the relationship between metacognition and learning at the level of individual knowledge units. One possibility is to increase the number of times at which PoP is measured within a single text. A problem with this solution is that frequently measuring PoP may interfere with text reading itself and thus reduce the ecological validity of the reading situation.

One should therefore look for procedures where metacognitive experiences can be measured more specifically in conjunction with different information units contained in the text, but where such measurement does not interfere with the reading process. One obvious candidate is the Judgment of Learning (JoL), which can be defined as "judgments made by participants at the end of a learning trial regarding the likelihood of remembering the acquired information on a subsequent memory test" (Koriat, 1997, p. 490). In other words, it refers to an item-specific prospective metamemory judgement (Metcalfe, 2000). What distinguishes it from other metamemory judgements (like for example Feelings of Knowing, Koriat, 1993) is that it is normally measured in the context of newly acquired knowledge rather than, for example, general semantic knowledge.

Although JoL, like PoP, is a measure of the predicted accuracy of future performance, it could be argued that JoL cannot straightforwardly be applied as a measure of metacognitive experiences in text reading. This is because, unlike PoP, it is rarely rated during the learning situation itself but most often at the end of the learning session, in conjunction with the presentation of a series of memory items. However, the focus in the present study is not so much on the phenomenology of metacognitive experiences during the learning process, as on metacognitive experiences related to the text material and their relationship to learning outcome. From this perspective, it could even be an advantage to measure metacognitive experiences after rather than during text reading. This is because learning outcome mainly reflects long-term memory, whereas a metacognitive rating given during reading mainly reflects short-term memory. This point was raised by Thiede, Griffin, Wiley, and Anderson (2010), as an argument against measuring metacognitive accuracy as the relationship between a metacognitive rating given during reading and performance on a subsequent memory task.

An alternative would be Confidence Ratings (CRs) conducted after participants have answered each of a series of recall/recognition questions (see, e.g., Norman & Price, 2015; for an introduction to CR measurement). In the context of memory, confidence refers to "the state of believing that a particular piece of information has been correctly retrieved from memory" (Miner & Reder, 1994). Because JoL is a prospective judgement related to a future event, whereas CR is a retrospective judgement related to a past event, JoL could be regarded as being more similar in phenomenology to PoP than is CR.

In this study we have included ratings of both JoL and CR, in addition to PoP. JoL and CR were measured on a trial-by-trial basis. To avoid the problem associated with few data points for the PoP judgement, whilst also controlling for which parts of the text each PoP was related to, we measured PoP twice for each text, i.e., halfway through each text and upon completion of each text.

1.2. Electronic reading devices

The only digital learning condition included in the studies by Ackerman and colleagues (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014) involved the use of a PC. In our view, it would also be interesting to include a condition in which reading occurs on different electronic reading devices (ERDs), because these are becoming more common, also as learning tools among students (Morineau, Blanche, Tobin, & Guéguen, 2005; Rockinson-Szapkiw, Courduff, Carter, & Bennett, 2013).

Existing studies looking at learning on ERDs have shown mixed results. According to some studies, learning is not influenced by whether a text is presented in the form of an e-book or a paper book (e.g., Rockinson-Szapkiw et al., 2013). Studies also report that the usefulness of ERDs stretches beyond that of PC's and paper books. For example, recent studies have reported that such devices "are a useful addition to laptops for the consumption of learning materials as well as for collaborative and social activities." (Fischer, Smolnik, & Galletta, 2013, p. 1). However some potential limitations have also been pointed out. For example, Morineau et al. (2005) argued that due to differences in the sensori-motor properties of e-books and paper books there is a risk that e-book presence "hinders recall of assimilated information whilst the presence of the paper support tends to facilitate it" (ibid, p.329). Moreover, Fischer et al. (2013) pointed out that various ERDs need further development in order to become useful for the production of content.

Even though existing research has looked at possible effects of ERDs on learning, no studies have specifically explored how it may affect the relationship between metacognitive experiences and learning outcome. It may be hypothesized that ERDs bridge the gap between on-screen and on-paper learning, in the sense that it shares some properties with the paper medium and others with the PC medium. When Ackerman and colleagues (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012) reported effects of study media on metacognitive calibration, with more overconfidence in a PC condition, the question arises as to whether the relationship between learning and metacognition in an ERD condition would most closely resemble the pattern observed in the paper or in the PC condition. It should be noted that different ERDs have different screen properties. For example, whereas the iPad-screen has a backlight that is comparable with a typical PC-screen, the Kindle screen has an E-ink display designed to resemble the visual properties of paper texts. Therefore, in Experiment 1 of the present study, we included two ERD study conditions (iPad and Kindle), in addition to a paper and a PC condition.

1.3. Individual differences in prior knowledge, interest, effort and strategy use

As pointed out by Ackerman and Goldsmith (2011), metacognitive calibration and resolution might be contaminated by factors other than on-line monitoring of metacognitive experiences, such as more stable individual differences in prior knowledge about the topic area. More studies are therefore needed which take into account individual difference variables that may be hypothesized to influence metacognitive experiences and/or learning outcome. In the most recent study, Lauterman and Ackerman (2014) included Scholastic Aptitude Test scores as a control variable. However, a number of other candidate variables that are likely to influence performance in a reading situation also need to be controlled for. These include the person's own estimate of their degree of prior knowledge, interest in the topic, and effort while reading the text.

An additional variable that may also influence learning and/or metacognition, and that may potentially also be influenced by study media, is learning strategies. Weinstein and Mayer (1986) define learning strategies as the behaviour and thoughts that take place during the learning process and that aims to promote learning. Learning strategies can be classified according to the depth of cognitive processing they involve, e.g., *superficial processing strategies* (e.g., memorization of information) or *deeper processing strategies* (e.g., organisation, elaboration and monitoring of information). The potential influence of information processing strategies on digital versus non-digital learning has been addressed by Liu (2005) and Morineau et al. (2005), who found that on-screen learning was associated with more shallow information processing, and on-paper learning with deeper information processing. As mentioned above, Ackerman and colleagues have repeatedly found that on-screen learners show more overconfidence than participants in an on-paper condition, with one potential explanation being that reading on screen may be associated with a more superficial approach. However strategy use was not directly measured in these studies. Therefore the current study included a self-report questionnaire (Anmarkrud & Bråten, 2009) that assessed these 2 forms of learning strategies in conjunction with text reading. We also included self-report measures of prior knowledge, interest in the topic, and effort while reading.

1.4. Note-taking and highlighting tools

When Ackerman and colleagues (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014) allowed their participants to take notes or mark the text while reading in order to control for the possible effect of note-taking on cognitive effort (Piolat, Olive, & Kellogg, 2005), they found that it did not interact with study media in the analyses of metacognitive resolution and calibration. However, we think that their testing and scoring procedure could be refined. First, although it was not stated explicitly in the articles, participants seem to have been given a score of one not only if they marked the text once, but also if they marked the text multiple times. The total score for each person therefore reflected only the number of texts for which note-taking and highlighting tools were used at all, and was not sensitive to differences in within-text note-taking and highlighting frequency among participants receiving a score above zero. Second, there was a difference in the number of note-taking and highlighting opportunities in the two conditions: The on-screen condition involved one extra highlighting tool (i.e., bold text) and one extra note-taking tool (i.e., inserting notes into the body text). Therefore, in Experiment 2 of the current study we gave the participants in the two conditions (PC and paper) an identical set of note-taking and highlighting tools (i.e., highlighting and commentary), and we calculated the absolute frequency, rather than the number of texts, with which each tool was used.

1.5. Research questions and hypotheses

In the current study, we conducted two experiments to explore the relationship between metacognitive experiences and recognition memory for on-paper versus on-screen learning by use of the measures described above. The difference between the two experiments were that Experiment 1 included 4 study conditions (paper, PC, iPad, Kindle) and did not include the opportunity for using note-taking and highlighting tools, whereas Experiment 2 included only 2 study conditions (paper, PC) but included the opportunity for using note-taking and highlighting tools. Such tools were not included in Experiment 1, partly due to the fact that ERDs are still relatively rare as learning tools in higher education, and that in a sample of university students some may have little experience with the use of ERD note-taking and highlighting tools.

The main research question was whether the relationship between learning outcome and metacognitive experiences would be influenced by study media. Because previous results are inconsistent with respect to whether learning outcome itself will be influenced by study media, this was treated as an open question. Regardless of possible differences in learning outcome across study conditions, we predicted that participants in the on-screen conditions would show more overconfidence than participants in the on-paper condition, as measured by the difference between PoP and learning outcome (i.e., metacognitive calibration), cf. the findings by Ackerman and colleagues (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014). Moreover, we predicted the moment-by moment relationship between each participant's PoP and their corresponding learning outcome (i.e., metacognitive resolution) to be similar for our PoP measure and the PoP measure used by Ackerman and Goldsmith (2011). However, we predicted on-paper learners to show higher metacognitive resolution than on-screen learners when the moment-by moment relationship between each participant's metacognitive score and their learning outcome was measured using JoL. This is because JoL measured in conjunction with individual knowledge units involves a higher degree of within-participant variance and therefore avoids the statistical limitations of PoP reported previously. Due to the fact that CR is temporarily further removed from the learning situation than JoL and PoP, in combination with a lack of relevant previous studies involving retrospective metacognitive judgements, we did not draw specific hypotheses regarding the influence of study media on the relationship between recognition performance and CR.

Based on the work of Liu (2005) and Morineau et al. (2005), we also predicted digital and non-digital learning to differ with respect to the involvement of shallow versus deep processing strategies.

Even though it can be hypothesized that electronic reading devices may bridge the gap between on-screen and on-paper learning, there are no previous studies of metacognitive monitoring for texts presented on ERDs in comparison to PC screens and paper. Therefore the pattern of results in this particular condition was treated as an open question. The same applied to possible differences between study conditions in the use of note-taking and highlighting tools, since previous studies including such variables have applied different procedures for measurement and scoring.

2. Experiment 1

2.1. Method

2.1.1. Participants

100 Norwegian students (52 males and 48 females) aged 18–32 ($M = 22.16$, $SD = 2.38$) were randomly assigned to one of 4 study media conditions, i.e., a paper condition, a PC condition, an iPad condition, or a Kindle condition, with 25 participants in each. Males and females were equally distributed between the four conditions. The participation was anonymous and each participant was rewarded NOK 100, − after the experiment.

2.1.2. Materials

The learning material consisted of 4 Norwegian texts of 1000 words each. Each text was 4 pages long, and contained 3 pictures. All texts were written in a 2-column layout, and were divided into shorter sections with subheaders.

In the 3 on-screen conditions participants read all text material in the form of a pdf document. In the paper condition, the same text material was printed in a booklet of A4-size paper. All other material was presented on paper for all participants, and participants always responded by pen and paper.

The titles of the texts were (translated from Norwegian) “Asfalt” (Text 1), “Tiled floors and electrostatic conductivity” (Text 2), “Hair transplantation” (Text 3), and “Molds in the production of cured meat” (Text 4). The texts were modified versions of texts found on different websites. The particular topics were chosen because we predicted that a student sample would have limited prior knowledge of them. All participants read the texts in the same order. For each text we developed 16 knowledge questions, e.g., “What was the name of the two authors who wrote a medical paper on hair transplantation in 1995?”, and a set of plausible response alternatives, e.g., “(a) Bernier and Russell, (b) Bernier and Rassmann, (c) Bernstein and Russell, (d) Bernstein and Rassmann”.

All measurements were organized into three questionnaires that will be explained in more detail in the procedure section.

2.1.3. Apparatus

In the PC condition the computer screen was 17” CRT monitor. The screen resolution was set to 1024 × 768. We used two different Ipad models (2nd and 3rd generation), with a 9.7” screen. For the Kindle condition we used Kindle DX, with a 9.7” screen.

The screen resolution for the PC condition was set so that the document size was the same as for the Ipad and Kindle conditions.

2.1.4. Procedure

Participants were tested in groups of 3–5 in separated cubicles in a purpose built psychology testing room. For reasons related to equipment and the counterbalancing of gender, sessions were mixed, i.e., different participants received different study media in the same session.

General instructions were given verbally and in writing at the start of the experiment. Participants were told that they would be presented with 4 texts that they were to read with the aim of learning the content, and that they would later be asked questions about this content. All participants in the PC condition reported being familiar with how to read a pdf document on a computer screen. The experimenter specifically instructed participants in the iPad and Kindle conditions on how to scroll and turn pages.

After being presented with the titles of the 4 texts, participants received *Questionnaire 1*, where they were first asked to indicate their background knowledge on each topic. The ratings were made on a 6-point scale from “very little” to “very much”. This variable will be referred to as *prior knowledge*. They were then given more specific information about the texts, i.e., that each text contained 1000 words, that they would be given 7 min to read each text, and that they would receive 16 questions related to each text. They were then asked to indicate how many of the 16 questions they thought they would be able to answer correctly. This variable will be referred to as *estimated ability*, and can be seen as an additional measure of prior knowledge.

Participants were then instructed to read the 4 texts. In the 3 on-screen conditions, the pdf document was open at the start of the experiment. The first page of the document contained a “Stop” sign with an instruction that one should not proceed to the next page (where Text 1 started) before instructed to do so. Participants were given 7 min to study each text. The experimenter indicated when they were allowed to start reading, and announced when 3.5 min had passed. *Questionnaire 2*, which contained PoP questions, questions about motivation and effort, as well as a learning strategy questionnaire, was administered together with the texts. For each text, Predictions of Performance (PoPs) were rated twice. At the bottom of page 2 of each text (after approximately 500 words), there was a text box instructing the participant to go to a particular question on a particular page of Questionnaire 2. They were here asked to indicate how many questions they thought they would be able to answer if given 8 questions concerning the piece of text they had just read. Upon completion of the text they were again instructed to go the questionnaire to give a second PoP rating. They were also asked to indicate how interesting they found the text, on a 6-point scale from “very uninteresting” to “very interesting”. This variable will be referred to as *interest*. They were then asked to indicate how much effort they had made while reading the text, on a 6 point scale from “very little” to “very much”. This variable will be referred to as *effort*. They were also asked to indicate whether they had managed to complete the whole text. After the participant had read all 4 texts, they answered a 24-item reading strategy questionnaire (Anmarkrud & Bråten, 2009). Each item was formulated as a statement, and one had to indicate, on a 6 point scale from “never” to “very often/always”, the extent to which the statement described what one had done while reading the 4 texts.

Questionnaire 3 contained the recognition task, JoL ratings and CR ratings. 64 factual questions (16 from each text) were presented twice, chronologically in the same order as the texts had been read. First, the questions were presented without any opportunity for participants to provide answers. Instead, for each question the participant had to indicate the likelihood that they would be able to recognize the correct answer to the question if presented among 4 alternatives. The scale was a 6-point scale from “very unlikely” to “very likely”. This rating is referred to as a Judgement of Learning (JoL) judgement. JoLs were self-paced. After making 64 JoL judgements participants were again presented with the same 64 questions in the same order, this time with 4 response alternatives for each question. The task was to select the correct answer and to indicate how certain one felt that the chosen answer was correct, on a 6-point scale from “very uncertain” to “very certain”. This rating is referred to as a Confidence Rating (CR). For both the JoL and recognition/CR tasks, participants were instructed to respond to the different items in the order they were presented, and they were not allowed to go back to previously completed items.

Parts of the study material and procedure has previously been applied in Furnes and Norman (2015).

2.2. Results

Relevant mean values and standard deviations are presented in Table 1.

2.2.1. Learning outcome for different texts

An ANOVA with learning outcome (i.e., number of correct recognition responses) as the dependent variable, and text (1–4) and study media (i.e, paper, PC, Ipad, Kindle) as independent variables, showed a significant main effect of text, $F(3,288) = 10.80$, $p < 0.001$, $\eta_p^2 = 0.10$, but no main effect of study media, $F(3,96) = 0.45$, $p = 0.72$, $\eta_p^2 = 0.01$, and no interaction between text and study media, $F(9,288) = 0.20$, $p = 0.99$, $\eta_p^2 = 0.01$. Because there was no interaction between text and study media, the following analyses are based on pooled data from all 4 texts.

Table 1
Mean scores and standard deviations on learning outcome, and various PoP, FoK, and CR scores in the four different study media, Experiment 1.

	Paper		PC		Ipad		Kindle	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Learning outcome	37.44	6.54	36.56	6.61	36.72	6.53	35.28	7.00
Total PoP	30.68	6.82	31.96	10.08	29.20	6.05	31.16	7.79
PoP calibration	−6.76	6.27	−4.60	11.68	−7.52	8.56	−4.12	8.05
PoP gamma	0.20	0.47	0.23	0.40	0.32	0.37	0.28	0.41
Mean JoL	3.91	0.58	3.99	0.66	3.91	0.57	3.94	0.62
JoL gamma	0.43	0.15	0.31	0.18	0.33	0.16	0.32	0.17
Mean CR	3.52	0.79	3.83	0.63	3.56	0.58	3.67	0.70
CR gamma	0.59	0.13	0.58	0.13	0.55	0.13	0.50	0.15
Deep strategies	3.04	0.87	3.17	0.62	3.17	0.59	3.15	0.79
Surface strategies	3.06	0.66	3.29	0.89	3.11	0.71	3.22	0.82
Effort	3.58	0.79	4.01	0.90	3.45	0.75	3.91	0.66
Prior knowledge	1.61	0.68	1.39	0.58	1.55	0.66	1.61	0.65
Estimated ability	8.71	1.88	7.84	3.51	8.45	2.09	8.46	1.95
Interest	2.94	0.91	3.24	0.78	2.72	0.76	3.15	0.76

2.2.2. Background variables

A series of ANOVAs showed that there were no significant group differences with respect to mean prior knowledge, $F(3,96) = 0.65$, $p = 0.58$, $\eta_p^2 = 0.02$, mean estimated ability, $F(3,96) = 0.57$, $p = 0.64$, $\eta_p^2 = 0.02$, or mean interest, $F(3,96) = 2.07$, $p = 0.11$, $\eta_p^2 = 0.06$. However there was a significant effect of study media on mean effort, $F(3,96) = 2.89$, $p = 0.04$, $\eta_p^2 = 0.08$. There was also a significant age difference between the groups, $F(3,96) = 5.65$, $p < 0.01$, $\eta_p^2 = 0.15$. A post-hoc analysis (Tukey's HSD test, all p's < 0.05) showed that this was due to the mean age being lower in the Kindle condition ($M = 20.64$, $SD = 1.66$) than in both the iPad condition ($M = 22.36$, $SD = 2.97$), the paper condition ($M = 22.56$, $SD = 2.31$), and the PC condition ($M = 23.08$, $SD = 1.71$). Thus, in the subsequent analyses, effort and age will be included as a covariate since it differed between study media.

2.2.3. Study media and learning outcome

An ANCOVA with learning outcome as the dependent variable, study media as the independent variable, and effort and age as covariates showed no significant effect of condition, $F(3,94) = 0.52$, $p = 0.67$, $\eta_p^2 = 0.02$.

2.2.4. Calculation and analysis of metacognition scores

Metacognitive experiences were assessed with Predictions of Performance (PoP), Judgements of Learning (JoL), and Confidence Ratings (CR). The main question of interest was the relationship between metacognition and learning, which was calculated for every participant as gamma correlations on the relationship between learning outcome and (a) PoP, (b) JoL, and (c) CR. In addition, we also calculated a PoP calibration bias score for each participant. For each form of metacognitive experience we also report the effect of study media on the rated metacognitive experience itself, even though this is not central to the main research question. All mean values and standard deviations are presented in Table 1.

2.2.5. Study media and PoP

An ANCOVA with mean PoP as the dependent variable, study media as the independent variable, and effort and age as covariates showed no significant effect of study media, $F(3,94) = 0.34$,

$p = 0.80$, $\eta_p^2 = 0.01$. For each participant we calculated a calibration bias score, i.e., the difference between the total PoP and total recognition score, with a positive score indicating overconfidence and a negative score indicating underconfidence. An ANCOVA with calibration scores as the dependent variable showed no effect of study media, $F(3,94) = 0.10$, $p = 0.96$, $\eta_p^2 < 01$. For each participant, we also conducted a gamma correlation between response accuracy and PoP. A subsequent ANCOVA with POP gamma score as the dependent variable showed no difference between conditions, $F(3,94) = 0.77$, $p = 0.51$, $\eta_p^2 = 0.02$.

2.2.6. Study media and JoL

An ANCOVA with mean JoL as the dependent variable, study media as the independent variable, and effort and age as covariates showed no significant effect of study media, $F(3,94) = 0.02$, $p = 0.99$, $\eta_p^2 < 0.001$. For each participant, we calculated a gamma correlation between response accuracy and JoL. An ANCOVA with this JoL gamma score as the dependent variable showed a significant effect of study media, $F(3,94) = 3.34$, $p = 0.02$, $\eta_p^2 = 0.10$. A post-hoc analysis (Tukey's HSD test) showed that this was due to the mean gamma score being significantly higher ($p = 0.04$) in the paper condition than in the PC condition.

2.2.7. Study media and CR

An ANCOVA with mean CR as the dependent variable, study media as the independent variable, and effort and age as covariates showed no effect of study media, $F(3,93) = 0.33$, $p = 0.80$, $\eta_p^2 = 0.01$. As for PoP and JoL, a gamma correlation was calculated between accuracy and CR for each participant. An ANCOVA with this CR gamma correlation as the dependent variable showed a non-significant trend for an effect of study media, $F(3,93) = 2.15$, $p = 0.10$, $\eta_p^2 = 0.06$.

2.2.8. Study media and strategy use

For each participant we calculated two strategy scores, one based on the seven items that comprise the *surface strategy* factor and the other based on the nine *deep strategy* items (Anmarkrud & Bråten, 2009). Two ANCOVAs were conducted with the respective strategy score factors as the dependent variable, study media as the independent variable, and effort and age as covariates. There were no significant effects of study media on either surface strategies, $F(3,94) = 0.27$, $p = 0.84$, $\eta_p^2 = 0.01$, or deep strategies, $F(3,94) = 0.18$, $p = 0.91$, $\eta_p^2 = 0.01$.

2.3. Discussion

There was no effect of study media on learning outcome. The relationship between metacognition and learning outcome did also not differ between groups when metacognitive experiences were measured as PoP or CR. However, a difference between conditions was identified for one metacognitive measure, namely JoL, suggesting that the relationship between JoL and learning outcome was stronger for participants in the paper condition compared to those in the PC condition. Contrary to predictions, there were no differences between the study media groups in self-reported learning strategies.

In order to test the replicability of the significant difference found with JoL, which was in contrast with the results involving PoP and CR, we conducted a second experiment which included only the two conditions that significantly differed from each other (i.e., paper and PC). In addition, to increase the ecological validity of the text reading situation, participants were given the opportunity to use note-taking and highlighting tools while reading.

3. Experiment 2

3.1. Method

3.1.1. Participants

50 Norwegian students (25 males and 25 females) aged 18–32 ($M = 22.60$, $SD = 3.00$) were randomly assigned to one of 2 study media conditions, i.e., a paper condition, and a PC condition, with 25 participants in each. Males and females were equally distributed between the two conditions. The participation was anonymous and each participant was rewarded NOK 100,- after the experiment.

3.1.2. Materials

The learning material and the questionnaire were identical to Experiment 1. Participants now read text material either on a pdf document (i.e., the PC condition) or in a printed A4 booklet (i.e., the paper condition).

3.1.3. Apparatus

The size and resolution of PC monitors were the same as in Experiment 1.

The difference from Experiment 1 was that participants were given the opportunity to highlight and make comments in the texts. Participants in the PC condition were allowed to use two markup tools in the Adobe Reader X, namely highlighting and sticky notes. Participants in the paper condition were provided with a yellow text marker and a ballpoint pen.

3.1.4. Procedure

The procedure was identical to that of Experiment 1, with the following two exceptions. Prior to reading the four texts, participants were informed of the opportunity to highlight and make comments in the texts. Participants in the PC condition were familiarized with the two markup tools. The experimenter first demonstrated how to use them. The participants were then instructed to try out the markup tools before starting to read Text 1. Because of this change in the procedure, we extended the reading time for each text from 7 to 12 min. The rationale for this extension was to give participants sufficient time to use these study tools in a way that more closely resembles a normal study situation, and thereby increase the ecological validity of the experiment.

3.2. Results

Relevant mean values and standard deviations are presented in Table 2.

3.2.1. Learning outcome for different texts

An ANOVA with learning outcome (i.e., number of correct recognition responses) as the dependent variable, and text (1–4) and study media (i.e, paper vs. PC) as independent variables showed no significant main effect of text, $F(3,144) = 2.29$, $p = 0.08$, $\eta_p^2 = 0.05$, no significant effect of study media, $F(3,144) = 3,54$, $p = 0.07$, $\eta_p^2 = 0.07$, and no significant interaction between text and study media, $F(3,144) = 1.65$, $p = 0.18$, $\eta_p^2 = 0.03$. As in Experiment 1, the subsequent analyses are based on pooled data from all 4 texts.

3.2.2. Background variables

A series of ANOVAs showed that there were no significant group differences with respect to mean prior knowledge, $t(48) = -0.07$, $p = 0.95$, $r = 0.10$, mean estimated ability, $t(48) = -1.62$, $p = 0.11$, $r = 0.23$, mean interest, $t(48) = 1.35$, $p = 0.18$, $r = 0.19$ or mean effort $t(48) = -1.31$, $p = 0.20$, $r = 0.19$. However, there was a significant age difference between the groups, $F(1,48) = 6.73$, $p = 0.01$, $\eta_p^2 = 0.12$, with a higher mean age in the paper condition

Table 2
Mean scores and standard deviations on learning outcome, and various PoP, JoL, and CR scores in the four different study media, Experiment 2.

	Paper		PC	
	M	*SD*	*M*	*SD*
Learning outcome	44.88	6.72	41.00	7.83
Total PoP	34.64	8.97	33.80	7.48
PoP calibration	−10.24	9.63	−7.20	8.79
PoP gamma	0.12	0.40	0.14	0.50
Mean JoL	4.60	0.57	4.24	0.57
JoL gamma	0.38	0.58	0.47	0.38
Mean CR	4.29	0.57	4.04	0.67
CR gamma	0.66	0.43	0.57	0.53
Deep strategies	3.47	0.61	3.63	1.14
Surface strategies	3.27	0.78	3.64	0.79
Effort	3.82	0.73	4.10	0.78
Prior knowledge	1.48	0.41	1.49	0.63
Estimated ability	8.08	2.91	9.27	2.25
Interest	3.58	1.05	3.22	0.83
Note-taking frequency	98.68	75.02	43.20	45.05

($M = 23.64$, $SD = 2.06$) than the PC condition ($M = 21.56$, $SD = 3.44$). Age was therefore included as a covariate in all the analyses.

3.2.3. Note-taking and highlighting

For each participant, we calculated the total number of comments and sentences that had been highlighted (partly or fully). For the PC condition this total note-taking score consisted of the number of times Adobe's "sticky note" function was used and the number of sentences that had been highlighted using Adobe's highlighting function. For the paper condition the note-taking score consisted of the number of individual words or clusters of words written in the margin of the text and the number of sentences that had been highlighted by the marker pen. There was a significant effect of study media on total note-taking frequency, $t(48) = 3.17$, $p = 0.003$, $r = 0.42$, with a higher mean score in the paper than in the PC condition.

In the following, we report a series of ANCOVAs with study media as the independent variable and note-taking frequency and age as covariates.

3.2.4. Study media and learning outcome

There was no effect of study media on learning outcome, $F(1,46) = 0.20$, $p = 0.65$, $\eta_p^2 < 0.01$.

3.2.5. Study media and PoP

There was no effect of study media on mean PoP, $F(1,46) = 0.08$, $p = 0.78$, $\eta_p^2 < 0.01$, on calibration bias scores, $F(1,46) = 0.36$, $p = 0.55$, $\eta_p^2 < 0.01$, or on the gamma correlation between response accuracy and PoP, $F(1,46) = 0.45$, $p = 0.50$, $\eta_p^2 = 0.01$.

3.2.6. Study media and JoL

There was no effect of study media on mean JoL, $F(1,46) = 2.14$, $p = 0.15$, $\eta_p^2 = 0.04$, or on the gamma correlation between response accuracy and JoL, $F(1,44) = 0.79$, $p = 0.38$, $\eta_p^2 = 0.02$. In other words, the effect of study media observed in Experiment 1 was not replicated under more ecologically valid study conditions and when controlling for note-taking frequency.

3.2.7. Study media and CR

There was no main effect of study media on mean CR, $F(1,46) = 1.69$, $p = 0.20$, $\eta_p^2 = 0.04$, or on the gamma correlation between response accuracy and CR, $F(1,42) = 0.04$, $p = 0.85$, $\eta_p^2 < 0.01$.

3.2.8. Study media and strategy use

There was a significant effect of study media on the surface strategy factor score, $F(1,46) = 5.65$, $p = 0.02$, $\eta_p^2 = 0.11$, with a higher mean score in the PC condition than in the paper condition. There was no effect of study media on the deep strategy factor score, $F(1,46) = 2.69$, $p = 0.11$, $\eta_p^2 = 0.06$.

3.3. Discussion

As in Experiment 1, the relationship between metacognition and learning outcome did not differ between groups when metacognitive experiences were measured as PoP or CR. However, we were not able to replicate the effect of study media on this relationship when metacognitive experiences were measured using JoL.

There was an effect of study media on the use of note-taking and highlighting tools, reflecting that participants in the paper condition used note-taking and highlighting tools more frequently than participants in the PC condition. When this measure was included as a covariate in the analysis comparing learning outcome across conditions, there was no significant difference between groups, which is compatible with the results of Experiment 1. However, without this measure as a covariate there was a nonsignificant trend for participants in the paper condition to show more learning. Therefore one cannot rule out the possibility that learning may have been influenced by the use of note-taking and highlighting tools.

For strategy use, there was a significant difference between groups, reflecting that participants in the PC condition reported more frequent use of surface strategies compared to the participants in the paper condition. This difference, however, was not found when comparing the use of deeper strategies across groups.

4. General discussion

In the present study, we conducted two experiments to investigate the relationship between metacognitive experiences and learning outcome across digital and non-digital learning contexts. More specifically, we examined the relationship between three different forms of metacognitive experiences and performance on a memory recognition task when the learning material was 4 different factual texts that the participant had read either on paper or on screen, where Experiment 1 included 3 on-screen conditions (PC, iPad, Kindle) and Experiment 2 included only 1 (PC). One measure of metacognitive experiences was PoP, a predictive judgement of future memory related to a particular part of a text, a measure previously applied in a similar context by Ackerman and colleagues (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014). However, to meet the procedural and statistical limitations associated with this measure we also included other measures that in our view are better suited for assessing metacognitive experiences on a moment-by-moment basis. These were JoL, i.e., a prospective memory judgement rated separately for each recognition item, and CR, i.e., a retrospective memory judgement rated separately for each recognition item. In both experiments we included prior knowledge, interest, and effort as background variables. We also measured self-reported reading strategies in conjunction with the various texts that the participants read. In Experiment 1 participants were not allowed to take notes and highlight text while reading. This opportunity was given to the participants in Experiment 2.

There was no indication in either experiment that metacognitive calibration, as measured by the absolute difference between PoP and subsequent recognition memory, was influenced by study media.

This is the metacognitive measure for which Ackerman and colleagues (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014) have most consistently found an effect of study media. More specifically, they found that learners in the on-screen condition were more overconfident than learners in the on-paper condition. The failure to replicate this finding goes against their hypothesis that on-screen learning is associated with less accurate self-regulation, which has been explained in terms of a variety of different factors. These include depth of information processing, which type of cues metacognitive judgements are based on, and the degree of match between study media and subjective media preference. It could be argued that our failure to replicate the effect repeatedly reported by Ackerman and colleagues might be due to the use of fixed study time, which was chosen to reduce confounding influences of, e.g., individual differences in reading speed, attention, and fatigue (Ackerman & Goldsmith, 2011). However, in previous studies the effect of study media has been at least as consistent with fixed as with free study time (see Ackerman & Goldsmith, 2011; Experiment 1; Ackerman & Lauterman, Experiments 1 and 2; Lauterman & Ackerman, 2014). On one occasion the effect was stronger with free study time (Ackerman & Goldsmith, 2011; Experiment 2), however on another it was stronger with fixed study time (Ackerman & Lauterman, 2012; Experiment 1). Therefore the use of fixed study time is unlikely to be the reason why we did not find any differences between the conditions in metacognitive calibration.

The absence of an effect of study media on metacognitive resolution as measured with PoP occurred in both experiments even though we had twice as many PoP measurements compared to previous studies (Ackerman & Goldsmith, 2011), and was not influenced by whether there was an opportunity to use note-taking and highlighting tools. In spite of the procedural modification compared to previous studies using PoP, this measure of metacognitive resolution was still based on a relatively low number of data points, which may reduce its sensitivity to pick up on genuine differences between the study media. Another and related point is that the relatively low number of measurements reduces the specificity of each PoP rating and thereby increases the likelihood that PoP ratings are instead influenced by factors unrelated to the learning situation itself, e.g., the individual's more global perceptions of their own abilities or their familiarity with the text topic. Ackerman and Goldsmith (2011) suggested that PoPs may sometimes be influenced by the individual's perception of their own general abilities. If so, PoPs may not be the most precise measure of metacognitive experiences in the context of text learning.

When metacognitive resolution was measured using JoL, a significant difference was found between the on-paper condition and the PC condition in Experiment 1, with no differences between either of these and the iPad and Kindle conditions. However, we were not able to replicate this finding in Experiment 2. This experiment specifically compared those two conditions (paper and PC) that differed in Experiment 1, in a reading situation that was made more realistic and ecologically valid by giving participants the opportunity to take notes and highlight the text while reading. A more cautious interpretation is therefore that the finding reported in Experiment 1 was spurious and indicative of a Type 1 error. Indeed, when a Bonferroni correction was applied, the result was no longer significant. Our JoL measurement was specifically designed to avoid the shortcomings associated with PoP, by allowing for a much larger number of data points in which each individual JoL judgement is made in conjunction with individual knowledge units. Since an effect of study media on metacognitive resolution was not found even when this highly specific measure of metacognitive experiences was applied, it seems unlikely that there are any identifiable differences in metacognitive resolution between study media. The fact that differences in metacognitive resolution were also not found across conditions when using CR, neither in Experiment 1 nor in Experiment 2, supports this interpretation. Another possibility is that the relationship between JoL and accuracy was influenced by the extent to which participants made use of the note-taking and highlighting tools. However this is unlikely to be the case since there was no difference in metacognitive resolution across the two conditions in Experiment 2 even when note-taking frequency was controlled for. Nevertheless, it might be that the opportunity to take notes influences cognitive processing in text reading independently of the frequency with which these tools are used, and that such an effect may mask effects of study media on metacognitive resolution that may be identified under other conditions. Regardless of which of these explanations are most probable, it seems that the significant influence of study media on metacognitive resolution that was observed in Experiment 1 for the JoL judgement is difficult to replicate under more ecologically valid study conditions. To summarize, there seems to be little evidence in support of the assumption that the relationship between metacognitive experiences and learning outcome is influenced by study media. Metacognitive accuracy was assessed by 4 analyses in each experiment, resulting in a total of 8 analyses, an effect of study media occurred in only one of these, but did not remain significant when controlling for multiple analyses. In addition, there was no effect of study media on either learning outcome or the mean value of any of the metacognitive measures, i.e., mean PoP, mean JoL, and mean CR, in either of the experiments. Moreover, there were no systematic differences in the use of deep and shallow strategies across study media. Together, the results indicate that at least when texts have a short, linear format, study media does not influence self-reported metacognitive experiences, learning outcome, or the relationship between the two.

5. Conclusions and implications

The results of the present study showed no systematic differences between digital and non-digital learning contexts in terms of the degree of correspondence between metacognitive monitoring and learning outcome. This contrast with previous studies that have claimed to demonstrate that on-screen learning is characterized by a lower correspondence between metacognitive experiences and learning outcome than on-paper learning (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012). Nevertheless, our study has some potentially important implications, both at a theoretical and an applied level.

The study presents a new procedure for comparing metacognitive monitoring for text reading across study media, and also contributes to basic research on JoL and CR, by exemplifying how these two may be measured in conjunction with the reading of short linear texts. Our results contribute to the understanding of the relationship between metacognitive experiences, i.e., a form of self-regulation, and learning in different study contexts.

The main applied contribution of our findings is that, at least for linear texts, there is little reason to believe that “screen inferiority” is related to possible differences in metacognitive experiences across study media. This is contrary to what the studies of Ackerman and colleagues have found. An implication of our findings is therefore that actions taken by individuals, educators and managers in order to adapt learning situations to digital study media and overcome “screen inferiority” should primarily focus on other mechanisms than metacognitive experiences.

Even though our results show little indication that study media influences metacognitive monitoring, it may still be the case that such an influence could be identified under other conditions. Therefore we want to suggest some directions for future research that may give us more insight into the relationship between metacognitive experiences and learning across different study media. First, as also pointed out by Ackerman and Goldsmith (2011), future studies should apply text material that is better adapted to a digital learning context, e.g., by using longer non-linear texts that includes hyperlinks, or by using multimodal texts where the message is conveyed through several information channels/modes, which are synchronized and integrated (Walsh, 2006). Second, one might consider using texts from other genres that may be associated with different patterns of reading strategies, e.g., fiction, newspaper articles, and instructional text. Finally, future studies should also consider increasing the total length of the text reading situation in the experiment and/or the time interval between text reading and memory retrieval. This might increase the ecological validity of the experiment, and thereby also the likelihood that any real-life differences in metacognitive monitoring across study media are detected.

Acknowledgements

This research was supported by a research grant to the first author from the Meltzer foundation (805440). We would also like to thank Professor Rune J. Krumsvik and the research group Digital Learning Communities at the University of Bergen, Norway, for supporting this project.

References

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: on screen versus on paper. *Journal of Experimental Psychology: Applied, 17*(1), 18–32. http://dx.doi.org/10.1037/a0022086.

Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior, 28*(5), 1816–1828. http://dx.doi.org/10.1016/j.chb.2012.04.023.

Anmarkrud, Ø., & Bråten, I. (2009). Motivation for reading comprehension. *Learning and Individual Differences, 19*, 252–256. http://dx.doi.org/10.1016/j.lindif.2008.09.002.

Benjamin, A. S., & Diaz, M. (2008). Measurement of relative metamnemonic accuracy. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of memory and metamemory* (pp. 73–93). New York, NY: Psychology Press.

Buzzetto-More, N., Sweat-Guy, R., & Elobaid, M. (2007). Reading in a digital age: e books: are students ready for this learning object? *Interdisciplinary Journal of Knowledge and Learning Objects, 3*, 239–250.

Carr, N. (2010). *The shallows: What the internet is doing to our brains*. New York, NY: N. N. Norton & Company.

Dascalua, M.-I., Bodea, C.-N., Lytras, M., Ordoñez de Pablos, P., & Burlacua, A. (2014). Improving e-learning communities through optimal composition of multidisciplinary learning groups. *Computers in Human Behaviour, 30*, 362–371. http://dx.doi.org/10.1016/j.chb.2013.01.022.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition* (1st ed.). London, UK: Sage Publications, Inc.

Fischer, N., Smolnik, S., & Galletta, D. F. (2013). Examining the potential for tablet use in a higher education context. In *Proceedings of the 11th International Conference on Wirtschaftsinformatik, 21 February–1 March, Leipzig, Germany* (pp. 9–22).

Furnes, B., & Norman, E. (2015). Metacognition and reading: Comparing three forms of metacognition in normally developing readers and readers with dyslexia. *Dyslexia, 21*, 273–284. http://dx.doi.org/10.1002/dys.1501.

Jamali, H. R., Nicholas, D., & Rowlands, I. (2009). Scholarly e-books: the views of 16,000 academics: results from the JISC national e-book observatory. *Aslib Proceedings, 61*(1), 33–47.

Koriat, A. (1993). How do we know that we know? the accessibility model of the feeling of knowing. *Psychological Review, 100*, 609–639. http://dx.doi.org/10.1037//0033-295X.100.4.609.

Koriat, A. (1997). Monitoring one's own knowledge during study: a cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*(4), 349–370. http://dx.doi.org/10.1037/0096-3445.126.4.349.

Koriat, A. (2000). The feeling of knowing: some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9*, 149–171. http://dx.doi.org/10.1006/ccog.2000.0433.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness* (pp. 289–325). New York: Cambridge University Press.

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior, 35*, 455–463. http://dx.doi.org/10.1016/j.chb.2014.02.046.

Lehmann, T., Hähnlein, I., & Ifenthaler, D. (2014). Cognitive, metacognitive and motivational perspectives on preflection in self-regulated online learning. *Computers in Human Behavior, 32*, 313–323. http://dx.doi.org/10.1016/j.chb.2013.07.051.

Liu, Z. (2005). Reading behavior in the digital environment: changes in reading behavior over the past ten years. *Journal of Documentation, 61*(6), 700–712. http://dx.doi.org/10.1108/00220410510632040.

Mangen, A. (2006). *New narrative pleasures? A cognitive-phenomenological study of the experience of reading digital narrative fictions* (Unplublished doctoral thesis). Trondheim, Norway: University of Science and Technology [NTNU]. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:no:ntnu:diva-1833.

Mangen, A., Waldermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: effects on reading comprehension. *International Journal of Educational Research, 58*, 61–68. http://dx.doi.org/10.1016/j.ijer.2012.12.002.

Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman–Kruskal gamma coefficient measure of association: implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 509–527. http://dx.doi.org/10.1037/a0014876.

Metcalfe, J. (2000). Metamemory: theory and data. In E. Tulving, & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 197–211). London: Oxford University Press.

Miner, A. C., & Reder, L. M. (1994). A new look at feeling of knowing: its metacognitive role in regulating question answering. In J. Metcalfe, & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 47–70). Cambridge: MIT Press.

Morineau, T., Blanche, C., Tobin, L., & Guéguen, N. (2005). The emergence of the contextual role of the e-book in cognitive processes through an ecological and functional analysis. *International Journal of Human-Computer Studies, 62*(3), 329–348. http://dx.doi.org/10.1016/j.ijhcs.2004.10.002.

Norman, E., & Price, M. C. (2015). Measuring consciousness with confidence ratings. In M. Overgaard (Ed.), *Behavioural methods in consciousness research* (pp. 159–180). Oxford, UK: Oxford University Press.

Pellas, N. (2014). The influence of computer self-efficacy, metacognitive self-regulation and self-esteem on student engagement in online learning programs: evidence from the virtual world of second life. *Computers in Human Behavior, 35*, 157–170. http://dx.doi.org/10.1016/j.chb.2014.02.048.

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*, 291–312. http://dx.doi.org/10.1002/acp.1086.

Rockinson-Szapkiw, A. J., Courdurff, J., Carter, K., & Bennett, D. (2013). Electronic versus traditional print textbooks: a comparison study on the influence of university students' learning. *Computers & Education, 63*, 259–266. http://dx.doi.org/10.1016/j.compedu.2012.11.022.

Spencer, C. (2006). Research on learners' preferences for reading from a printed text or from a computer screen. *Journal of Distance Education, 21*(1), 33–50.

Thiede, K. W., Griffin, T. D., Wiley, J., & Anderson, M. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. *Discourse Processes, 47*, 331–362. http://dx.doi.org/10.1080/01638530902959927.

Walsh, M. (2006). The "textual shift": examining the reading process with print, visual and multimodal texts. *Australian Journal of Language and Literacy, 29*(1), 24–37.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. *Handbook of Research on Teaching, 3*, 315–327.

Woody, W. D., Daniel, D. B., & Baker, C. A. (2010). E-books or textbooks: students prefer textbooks. *Computers & Education, 55*(3), 945–948. http://dx.doi.org/10.1016/j.compedu.2010.04.005.

Zhao, J., & Ordóñez de Pablos, P. (2011). Regional knowledge management: the perspective of management theory. *Behaviour & Information Technology, 30*(1), 39–49. http://dx.doi.org/10.1080/0144929X.2010.492240.

Zimmerman, B. J. (1986). Becoming a self-regulated learner: which are the key subprocesses? *Contemporary Educational Psychology, 11*(4), 307–313. http://dx.doi.org/10.1016/0361-476X(86)90027-5.